



Ottawa Hull K1A 0C9

(21) (A1) 2,108,113
(22) 1993/10/08
(43) 1995/04/09

BEST AVAILABLE COPY

(51) INTL.CL.⁵ C12N-015/52; C12N-015/76; C12N-009/00; C12P-017/18;
C12N-001/21

(19) (CA) **APPLICATION FOR CANADIAN PATENT** (12)

(54) DNA Sequence Encoding Enzymes of Clavulanic Acid
Biosynthesis

(72) Jensen, Susan E. - Canada ;
Aidoo, Kwamena A. - Canada ;
Paradkar, Ashish S. - Canada ;

(71) Governors of the University of Alberta (The) - Canada ;

(57) 39 Claims

5,095,2/24

Notice: This application is as filed and may therefore contain an
incomplete specification.



Industrie Canada Industry Canada

3488

Canada

ABSTRACT

DNA sequences are provided which encode the enzymes required for clavulanic acid synthesis. A process is
5 provided for producing clavulanic acid in a transformant of a non-clavulanate-producing host.

DNA SEQUENCE ENCODING ENZYMES OF CLAVULANIC ACID
BIOSYNTHESIS

This invention relates to methods for the production
5 of the antibiotic, clavulanic acid.

Background of the Invention

Clavulanic acid is a broad spectrum beta-lactamase
inhibitor and is an important antibiotic for the
10 treatment of infectious diseases. It is produced
commercially by the gram-positive mycelial prokaryote
Streptomyces clavuligerus, which also produces the β -
lactam antibiotics penicillin N, desacetox
cephalosporin C and cephamycin C. Until recently,
15 however, the pathway employed for clavulanic acid
biosynthesis was much less well understood than the
pathways leading to these other antibiotics.

Without knowledge of the pathway for clavulanic acid
biosynthesis, it was not possible to isolate the genes
20 coding for the key enzymes and to manipulate these genes
to increase antibiotic yield or permit production of the
antibiotic in heterologous systems.

One of the earliest enzymes of the pathway to be
purified and characterised was clavaminic acid synthase.
25 Two isozymes have now been identified and characterised
(Marsh et al., (1992), Biochem., vol. 31, pp. 12648-657).

European Patent Application 0349121 describes a DNA
restriction fragment encoding a portion of the genetic
information involved in clavulanic acid synthesis but
30 provides no sequence information.

Until the work of the present inventors, the
complete complement of genes required for clavulanic acid
synthesis had not been identified. The present inventors
have now isolated, cloned and sequenced an 11.6 kb
35 genomic DNA sequence from S. clavuligerus which codes for
eight proteins and enables the production of clavulanic

acid by transformants of non-clavulanic-producing organisms.

Summary of the Invention

- 5 An isolated genomic DNA molecule is provided comprising the nucleotide sequence set out in Figure 2. A process is provided for producing clavulanic acid in a transformant of a non-clavulanate-producing host.

10 Description of Drawings

The invention, as exemplified by a preferred embodiment, is described with reference to the accompanying drawings in which:

- 15 Figure 1 shows the N terminal amino acid sequence of CLA and the nucleotide sequence of a probe (Sequence ID No.:2) directed to the underlined region of the sequence.

- Figure 2 (2-1 to 2-10) shows the nucleotide sequence (Sequence ID No.:1) of a 15 kb genomic DNA fragment from S. clavuligerus. The sequences of the ten ORFs within
20 the fragment are shown in upper case letters and the intergenic regions are shown in lower case letters. The locations of the beginning and end of each ORF are also indicated directly above the nucleotide sequence.
25 Asterisks above the sequence indicate the EcoRI sites which mark the beginning and end of the portion of the DNA sequence which contains all the genetic information for clavulanic acid synthesis.

 Figure 3 shows the location of the open reading frames downstream from pcbC.

- 30 Figure 4 shows a partial restriction map of the DNA sequence of Figure 2 in the region surrounding cla (ORF4).

 Figure 5 shows a shuttle vector used for disruption of the cla gene.

- 35 Figure 6 shows a photograph of an agar plate bearing cultures of S. lividans transformants.

Figure 7 shows an alignment of the amino acid sequence of CLA (S. clavuligerus CLA) with those of E. Coli agmatine ureohydrolase (E. Coli AUH), yeast arginase (yeast ARG), rat arginase (rat ARG) and human arginase (human ARG).

Figure 8 shows a Southern blot of NcoI digests of genomic DNA from five presumptive mutants (lanes 1-5) and from wild-type S. clavuligerus (lane 6). Panel A : membranes probed with cla-specific probe. Panel B : membranes probed with tsr-specific probe.

Figure 9 shows restriction enzyme maps of S. clavuligerus DNA inserts in cosmids. A. Restriction enzyme map of cosmid K6L2. B. Partial restriction enzyme map of cosmid K8L2. C. Restriction map of cosmids K6L2 and K8L2 indicating location of pcbC gene in relation to cla. D. The 2.0 kb NcoI fragment encompassing the cla gene used in generating nested deletions for sequencing. Abbreviations: Ba, BamHI; B, BglIII; E, EcoR1; K, KpnI; N, NcoI; S, SalI; and Sm, SmaI.

Figure 10 shows the deduced amino acid sequence (Sequence ID No.:3) of ORF1 of Figure 2.

Figure 11 shows the deduced amino acid sequence (Sequence ID No.:4) of ORF2 of Figure 2.

Figure 12 shows the deduced amino acid sequence (Sequence ID No.:5) of ORF3 of Figure 2.

Figure 13 shows the deduced amino acid sequence (Sequence ID No.:6) of ORF4 of Figure 2.

Figure 14 shows the deduced amino acid sequence (Sequence ID No.:7) of ORF5 of Figure 2.

Figure 15 shows the deduced amino acid sequence (Sequence ID No.:8) of ORF6 of Figure 2.

Figure 16 shows the deduced amino acid sequence (Sequence ID No.:9) of ORF7 of Figure 2.

Figure 17 shows the deduced amino acid sequence (Sequence ID No.:10) of ORF8 of Figure 2.

Figure 18 shows the deduced amino acid sequence (Sequence ID No.:11) of ORF9 of Figure 2.

Figure 19 shows the deduced amino acid sequence (Sequence ID No.:12) of ORF10 of Figure 2.

Detailed description of the Invention

5 Production of penicillin and cephamycin antibiotics in S. clavuligerus starts with the conversion of lysine to α -aminoadipic acid (Madduri et al., (1989), J. Bacteriol., v. 171, pp. 299-302; (1991), J. Bacteriol., v. 173, pp. 985-988). α -Aminoadipic acid then condenses
10 with cysteine and valine to give δ -(L- α -aminoadipyl)-L-cysteinyl-D-valine (ACV) by the action of aminoadipyl-cysteinyl-valine synthetase (ACVS). ACV is converted by isopenicillin N synthase (IPNS) to isopenicillin N, and, through a series of reactions, to desacetoxycephalosporin
15 C and ultimately to cephamycin C (Jensen et al., (1984), Appl. Microbiol. Biotechnol., v. 20, pp 155-160).

 The ACVS of S. clavuligerus has been purified and partially characterized by three separate groups, and estimates of its molecular weight vary from 350,000 to
20 500,000 Da (Jensen et al., (1990) J. Bacteriol., v. 172, pp. 7269-7271; Schwecke et al., (1992); Eur. J. Biochem., v. 205, pp. 687-694; Zhang and Demain, (1990), Biotech Lett., v. 12, pp. 649-654). During their purification, Jensen et al. observed a 32,000 Da protein which co-
25 purified with ACVS despite procedures which should remove small molecular weight components. It has now been found that this protein is not related to ACVS but rather to clavulanic acid biosynthesis. It has been designated CLA.

30 In accordance with one embodiment of the invention, the present inventors have identified, cloned and sequenced the gene (cla) encoding this protein.

 In accordance with a further embodiment of the invention, the inventors have cloned and sequenced a 15
35 kb stretch of genomic DNA from S. clavuligerus which includes the cla gene. Within this 15 kb sequence, the inventors have identified an 11.6 kb DNA fragment which,

when introduced into the non-clavulanate producer S. lividans as described in Example 4, enabled that species to produce clavulanic acid. This indicates that the 11.6 kb fragment contains all the genetic information required for clavulanate production.

As will be understood by those skilled in the art, the identification of the DNA sequence encoding the enzymes required for clavulanate synthesis will permit genetic manipulations to modify or enhance clavulanate production. For example, clavulanate production by S. clavuligerus may be modified by introduction of extra copies of the gene or genes for rate limiting enzymes or by alteration of the regulatory components controlling expression of the genes for the clavulanate pathway.

Heterologous organisms which do not normally produce clavulanate may also be enabled to produce clavulanate by introduction, for example, of the 11.6 kb DNA sequence of the invention by techniques which are well known in the art, as exemplified herein by the production of S. lividans strains capable of clavulanate synthesis. Such heterologous production of clavulanic acid provides a means of producing clavulanic acid free of other contaminating clavams which are produced by S. clavuligerus.

Suitable vectors and hosts will be known to those skilled in the art; suitable vectors include pIJ702, pJOE829 and pIJ922 and suitable hosts include S. lividans, S. parvulus, S. griseofulvus, S. antibioticus and S. lipmanii.

Additionally, the DNA sequences of the invention enable the production of one or more of the enzymes of the clavulanate pathway by expression of the relevant gene or genes in a heterologous expression system.

The DNA sequences coding for one or more of the pathway enzymes may be introduced into suitable vectors and hosts by conventional techniques known to those skilled in the art. Suitable vectors include pUC118/119

and pET-11 and suitable hosts include many organisms, including E. coli strains such as MV1193 and BL21(DE3).

An oligonucleotide probe based on the N-terminal amino acid sequence of CLA was constructed as shown in Figure 1 and was used to isolate the gene coding for the protein from S. clavuligerus, as described in Example 1.

The gene was found to be located in the S. clavuligerus chromosome about 5.7 kb downstream of pcbC, the gene which encodes isopenicillin N synthase. The gene contains a 933 bp open reading frame (ORF), encoding a protein of molecular weight 33,368. The deduced amino acid sequence was compared to database sequences and showed greatest similarity to enzymes associated with arginine metabolism, notably agmatine, ureohydrolase and arginases.

When an internal fragment of the cla gene was labelled and used to probe restriction endonuclease digests of genomic DNA from a variety of other Streptomyces and related species, evidence of homologous sequences was seen only in other clavulanic acid or clavam metabolite producers, including Streptomyces jumonjinensis, Streptomyces lipmanii (7) and Streptomyces antibioticus. No cross reactivity was seen to the β -lactam producing species Nocardia lactamdurans, Streptomyces griseus or Streptomyces cattleya, nor to any of a variety of other Streptomyces species which do not produce β -lactam compounds, including S. fradiae ATCC 19609, S. venezuelae 13s and S. griseofulvus NRRL B-5429.

Disruption of the cla gene, as described in Example 3, led to loss of the ability to synthesise clavulanic acid.

A 15 kb DNA sequence extending downstream from pcbC was cloned and sequenced as described in Example 5. The nucleotide sequence is shown in Figure 2. When this sequence information was analysed for percent G + C as a function of codon position (Bibb et al., (1984), Gene, v. 30, pp. 157-166), ten complete ORFs were evident, as

shown in Figure 3. ORF 4 corresponds to cla. ORF 1, 7 & 8 are oriented in the opposite direction to pcbC. ORFs 2-6 and ORF 10 are all oriented in the same direction as pcbC. ORFs 2 and 3, and ORFs 4 and 5 are separated by very short intergenic regions suggesting the possibility of transcriptional and translational coupling. Table 1 summarises the nucleotide sequences and lengths of ORFs 1-10.

When the predicted amino acid sequences of proteins encoded by ORFs 1 - 10 were compared to protein sequence databases, some similarities were noted in addition to the already mentioned similarity between CLA and enzymes of arginine metabolism. ORF 1 showed a low level of similarity to penicillin binding proteins from several different microorganisms which are notable for their resistance to β -lactam compounds.

An EcoRI fragment of the 15 kb DNA sequence, containing 11.6 kb DNA, was cloned into a high copy number shuttle vector and introduced into S. lividans, as described in Example 4. Of seventeen transformants examined, two were able to produce clavulanic acid, indicating that the 11.6 kb fragment contains all the necessary genetic information for clavulanic acid production.

This 11.6 kb fragment encompasses ORF 2 to ORF 9 of the 15 kb DNA sequence.

ORF 2 shows a high degree of similarity to acetohydroxyacid synthase (AHAS) enzymes from various sources. AHAS catalyses an essential step in the biosynthesis of branched chain amino acids. Since valine is a precursor of penicillin and cephamycin antibiotics, and valine production is often subject to feedback regulation, it is possible that a deregulated form of AHAS is produced to provide valine during the antibiotic production phase. Alternatively, an AHAS-like activity may be involved in clavulanic acid production. While the presently recognized intermediates in the clavulanic acid

biosynthetic pathway do not indicate a role for AHAS, the final step in the biosynthetic pathway, conversion of clavaminic acid to clavulanic acid, requires NADPH, and either pyruvate or α -ketobutyrate as well as other cofactors (Elson et al., (1987), J. Chem. Soc. Chem. Commun., pp. 1739-1740). It is striking that these same substrates and cofactors are required for AHAS activity. Perhaps the conversion of clavamate to clavulanate actually involves several steps, one of which is catalyzed by an AHAS-like activity. ORFs 3 and 5 do not show a significant similarity to any proteins in the data bases. ORF 6 shows similarity to ornithine acetyltransferase. Ornithine has been suggested to be the immediate precursor of a 5-C fragment of the clavulanic acid skeleton, but the details of the reaction required for the incorporation of ornithine are unknown. ORF 7 shows weak similarity to protein XP55 from S. lividans, and a lower level of similarity to oligopeptide binding proteins from various other species. Similarly, ORF 8 shows weak similarity to several transcription activator proteins, and ORF 9 shows weak similarity to ribitol 5 PO₄ dehydrogenase-type enzymes. ORF 10 shows a high similarity to cytochrome P450 type enzymes from other Streptomyces species.

ORF5 has now been identified as the gene for clavamate synthase II (Marsh (1993) supra).

When a plasmid isolated from one of the two clavulanic acid-producing transformants was retransformed into S. lividans, about 40-45% of the resulting colonies were able to produce clavulanic acid, as shown in Figure 6.

EXAMPLESExample 1Bacterial strains, vectors and growth conditions.

- 5 Streptomyces clavuligerus NRRL 3585, Streptomyces
 jumoniensis NRRL 5741, Streptomyces lipmanii
 NRRL 3584, Streptomyces griseus NRRL 3851, Nocardia
 lactamdurans NRRL 3802 and Streptomyces cattleya NRRL
10 3841 were provided by the Northern Regional Research
 Laboratories, Peoria, IL. Streptomyces antibioticus ATCC
 8663 and Streptomyces fradiae ATCC 19609 were obtained
 from the American Type Culture Collection, Rockville, MD.
 Streptomyces lividans strains 1326 and TK24 were provided
15 by D.A. Hopwood (John Innes Institute, Norwich, U.K.),
 Streptomyces venezuelae 13s and Streptomyces griseofuscus
 NRRL B-5429 were obtained from L.C. Vining (Department of
 Biology, Dalhousie University, Halifax, N.S.). Cultures
 were maintained on either MYM (Stuttard (1982) J. Gen.
 Microbiol., v. 128, pp. 115-121) or on a modified R5
20 medium (Hopwood et al. (1985) in "Genetic Manipulation of
 Streptomyces : a laboratory manual", John Innes
 Foundation, U.K.) containing maltose instead of glucose
 and lacking sucrose (R5-S). Escherichia coli MV1193
 (Zoller and Smith (1987) Methods in Enzymology, v. 154,
25 pp. 329-349), used as recipient for all of the cloning
 and subcloning experiments, was grown in Luria Broth (LB;
 Sambrook et al. (1989) in "Molecular Cloning : a
 laboratory manual", Cold Spring Harbour, N.Y.) or on LB
30 agar (1.5%) plates containing ampicillin (50 µg/mL) or
 tetracycline (10 µg/mL). The cloning vectors pUC118 and
 pUC119 (Vieira and Messing (1987) Methods in Enzymology,
 v. 153, pp. 3-11) were provided by J. Vieira (Waksman
 Institute of Microbiology, Rutgers University,
 Piscataway, N.J.). The plasmid vector pJOE829 was
35 generously provided by J. Altenbuchner (University of
 Stuttgart, Stuttgart, Germany). The plasmid pIJ702 was
 obtained from the American Type Culture Collection,

Rockville, MD. Restriction enzymes were purchased from Boehringer Mannheim, and used according to the manufacturers' specifications.

5 Separation of CLA from ACVS

CLA was previously characterized as a 32,000 Da molecular weight protein present in preparations of highly purified ACVS (Jensen et al. (1990), supra). The small size of CLA suggested that its co-purification with
10 ACVS resulted from a physical association between the two proteins.

ACVS and CLA were resolved by applying a 0.2 ml sample of purified ACVS containing CLA onto a Superose 6 HR 10/30 (Pharmacia), which was equilibrated and eluted
15 in 0.1 M MOPS buffer, pH 7.5 containing 0.05 M KCl, 1 mM dithiothreitol, and 20% glycerol, at a flow rate of 0.25 ml/min.

Comparison of the CLA retention time with those of molecular weight standards indicated that the native
20 molecular weight of CLA was in excess of 270 kDa. The difference in molecular weight between native and denatured forms of CLA suggests that the native protein exists as an oligomer of eight identical subunits.

25 Isolation of gene (cla) for CLA

N-terminal amino acid sequence information for CLA was obtained by electrophoretically transferring the protein from SDS polyacrylamide gels onto Immobilon membranes (Millipore Ltd.,) and submitting the material
30 to the Protein Microsequencing Laboratory (University of Victoria,) for analysis. Information obtained for 25 amino acids at the N-terminus was used to prepare a 24-mer oligonucleotide probe with 8-fold degeneracy to the amino acid sequence underlined in Figure 1. The amino
35 acids in brackets indicate ambiguities in the N-terminal sequence. The actual DNA sequence from the cloned fragment is indicated in Figure 1.

The probe was designed as an 8-fold degenerate mixture of oligonucleotides to take into consideration the biased codon usage of Streptomyces (Bibb et al., 1984, Wright and Bibb (1992), Gene, v. 113, pp. 55-65).).

5 End-labelled probe was then used to screen a cosmid library of S. clavuligerus genomic DNA fragments as described in Materials and Methods.

A library of S. clavuligerus genomic DNA fragments (15-22 kb size fractionated fragments) was constructed as
10 previously described (Doran et al. (1990), J. Bacteriol., v. 172, pp. 4909-4918). using the cosmid vector pLAFR3. A collection of 1084 isolated E. coli colonies containing recombinant cosmids was screened for the presence of cla using the 24-mer mixed oligonucleotide probe (Fig. 1)
15 which had been end-labelled with [γ -³²P]dATP and polynucleotide kinase (Boehringer Mannheim). Colony hybridization and subsequent washing was performed as described by Sambrook et al., (1989), at 55°C with a final wash in 0.2X SSC (1X SSC, 0.15M NaCl and 0.015M
20 sodium citrate) and 0.1% SDS.

Five colonies which gave strong hybridization signals were isolated from the panel of 1084 clones, and restriction analysis showed that the positive clones contained overlapping fragments of DNA. Two clones, K6L2
25 and K8L2, with sequences that spanned about 40 kb of the S. clavuligerus genome, were chosen for further analysis. Clone K8L2 contained about 22 kb of S. clavuligerus genomic DNA and included a portion of cla and all of the pcbC gene which encodes IPNS in the penicillin/cephamycin
30 biosynthetic pathway. A restriction map of K6L2 is shown in Fig. 9. Within the approximately 27 kb of DNA contained in K6L2, the oligonucleotide probe hybridized to a 2.0 kb NcoI fragment which was subsequently found to contain the entire cla gene. Hybridization studies,
35 restriction mapping and DNA sequence analysis revealed that cla was situated 5.67 kb downstream of the pcbC gene of S. clavuligerus (Fig. 9).

DNA sequencing and analysis

Ordered sets of deletions were generated (Henikoff, 1984) extending across the cla region of the 2.0 kb NcoI fragment (Fig. 9C). The deletion generated fragments
5 were sequenced in both orientations by the dideoxynucleotide chain termination method of (Sanger et al. (1977), P.N.A.S., v. 74, pp. 5463-5467) using Sequenase (version 2.0) DNA polymerase (United States Biochemical Corporation). Areas of compression in the
10 sequence band pattern were relieved by carrying out reactions using 7-deaza-dGTP in place of dGTP. The nested deletion fragments resided either in pUC118 or pUC119, and were sequenced using the commercially available universal primers (Vieira and Messing, 1987).

15 The nucleotide sequence data were analyzed for the presence of restriction sites, open reading frames (ORFs) and codon usage by the PC-Gene programme (Intelligenetics Corp.). Similarity searches were accomplished with the FASTA program searching the GenPept database (release
20 number 71) available through GenBank (Pearson and Lipman (1988), P.N.A.S., v. 85, pp. 2444-2448).

An ORF of 939 bp with a potential ribosome site 9 bp from the GTG start codon was found which encoded a putative protein with a molecular weight of 33,368 Da.
25 This value is in close agreement to the molecular weight estimated for CLA by SDS-PAGE (Jensen et al., 1990). The analysis of percent G + C as a function of codon position (FRAME analysis), using the algorithm of Bibb et al., (1984), indicated the presence of a typical streptomycete
30 ORF (data not shown) with a G + C content of 70%. Computer aided data base searches for sequences similar to cla revealed a high degree of similarity to agmatine ureohydrolase (40.5% identity over 291 amino acids) and somewhat lower similarity to arginases (29.6% identity
35 over 135 amino acids to arginases from yeast and rat) as shown in Figure 7. The S. clavuligerus CLA sequence was aligned with the E. coli AUH sequence by the FASTA

program described above. The AUH sequence had previously been aligned with the three ARG sequences (Szumanski & Boyle (1990), J. Bacteriol., v. 172, pp. 538-547).

Identical matches in two or more sequences are indicated with upper case letters.

Example 2

DNA hybridization

Genomic DNA preparations from various Streptomyces species were isolated as described by Hopwood et al. (1985). For interspecies DNA hybridization analysis, 2.0 μ g amounts of genomic DNA preparations were digested with NcoI for 16h, and electrophoresed in 1.0% agarose gels. The separated DNA fragments were then transferred onto nylon membranes (Hybond-N, Amersham) and hybridized with a cla specific probe prepared by labeling an internal 459 bp SalI fragment (Fig. 1) with [α - 32 P]dATP by nick translation. Hybridization was done as described by Sambrook et al., (1989). Hybridization membranes were washed twice for 30 min in 2X SSC; 0.1% SDS and once for 30 min in 0.1X SSC; 0.1% SDS at 65°C.

Sequences homologous to cla in other Streptomycetes

Three of six producers of β -lactam antibiotics, S. clavuligerus, S. lipmanii and S. jumonjinensis showed positive hybridization signals whereas S. cattleya, S. griseus, and N. lactamdurans did not (data not shown). None of the nonproducing strains examined, S. venezuelae, S. lividans, S. fradiae, S. antibioticus and S. griseofuscus gave any signal. All of the streptomycetes that gave positive signals were producers of clavam-type metabolites (Elson et al., 1987)

Example 3

Disruption of the genomic cla gene

A 2.0 kb NcoI fragment that contained the entire cla gene was digested at its unique KpnI site and the ends

made blunt by treatment with the Klenow fragment of E. coli DNA polymerase I. A thiostrepton resistance gene (tsr), isolated as a 1085 bp BclI fragment from pIJ702 and cloned into the BamHI site of pUC118 was excised as a
5 SmaI/XbaI fragment and the ends made blunt as above and ligated into the KpnI site of cla. The ligation mixture was introduced into E. coli MV1193 and the transformants screened for the presence of the tsr gene by colony hybridization (Sambrook et al., 1989).

10 Replacement of the chromosomal cla gene by a copy disrupted by the insertion of tsr, at an internal KpnI site, was achieved by double recombination. Successful gene replacement was apparent when the 2.0 kb NcoI
15 fragment which carries cla in the wild type organism was replaced by a 3.0 kb NcoI fragment due to the insertion of the 1.0 kb tsr gene in the mutants. Four of the five mutants tested showed the expected increase in the size of the NcoI fragments, and the larger NcoI fragments also hybridized with a tsr specific probe. The fifth mutant
20 was apparently a spontaneous theostrepton resistant mutant.

Antibiotic Assay

The agar diffusion assay was used for determining
25 both penicillin/cephamycin and clavulanic acid production. S. clavuligerus strains to be assayed were grown in 10 ml. amounts of Trypticase Soy Broth (TSB; Baltimore Biological Laboratories) medium with 1.0% starch for 48h. The cultures were washed twice with
30 10.3% sucrose and once with MM (Jensen et al. (1982), J. Antibiot., v. 35, pp. 483-490) and the mycelium resuspended in 10.0 mL of MM. Two millilitres of washed cell suspension was inoculated into 100 mL of MM and incubated at 28°C for 48h. The cultures were harvested
35 by centrifugation, and the supernatants were assayed for both penicillin/cephamycin and clavulanic acid using

bioassay procedures described previously (Jensen et al. (1982), supra).

All of the resulting colonies with disrupted cla genes grew equally well on minimal medium and complex media and produced as much penicillin and cephamycin as did the wild-type, but produced no clavulanic acid (data not shown). HPLC analysis of cell supernatants confirmed the inability of the disrupted cla mutants to synthesize any clavulanic acid (data not shown).

10

Example 4

Protoplast formation and transformation

E. coli competent cell preparation and transformation were as described by Sambrook et al., (1989). Protoplasts of S. clavuligerus were, prepared, transformed and regenerated as described by Bailey et al. (1984), Bio/Technology, v. 2, pp. 808-811, with the following modifications. Dextrin and arginine in the regeneration medium were replaced by starch and sodium glutamate respectively. Protoplasts were heat shocked at 43°C for 5 min prior to the addition of DNA. Standard procedures were used for protoplasting and transformation of S. lividans (Hopwood et al. (1985)).

The 11.6 kb EcoRI fragment from K6L2 (Fig. 9) was cloned into the EcoRI site of pCAT-119. pCAT-119 is derivative of pUC119 which was prepared by insertionally inactivating the ampicillin resistance gene of pUC119 by the insertion of a chloramphenicol acetyltransferase gene (Jensen et al. (1989), Genetics & Molec. Biol. of Ind. Microorg., pp. 239-245 Ed. Hershberger, Amer. Soc. Microbiol.). The PCAT-119 plasmid carrying the 11.6 kb fragment was then digested with PstI and ligated to the Streptomyces plasmid pIJ702, which had also been digested with PstI. The resulting bifunctional plasmid carrying the 11.6kb insert was capable of replicating in either E. coli (with selection for chloramphenicol resistance) or in S. lividans (with selection for thiostrepton

30

35

resistance). The ligation mixture was transformed to E. coli. Plasmid DNA was isolated from several of the chloramphenicol resistant transformants and analyzed by agarose gel electrophoresis to ensure that the proper plasmid construct was obtained. This isolated plasmid material from E. coli was then transformed into S. lividans as described by Hopwood and transformants were selected by plating onto R2YE medium containing thiostrepton at a concentration of 50 µg/ml.

Thiostrepton resistant S. lividans transformants carrying the bifunctional plasmid with the 11.6 kb insert were patched onto MYM agar plates and allowed to incubate for 48h at 28°C before they were overlaid with molten soft nutrient agar containing penicillin G at a concentration of 1 µg/ml and inoculated with Staphylococcus aureus N-2 as indicator organism (Jensen, 1982). (S. aureus N-2 was obtained from the Department of Microbiology Culture Collection, University of Alberta. Any organism which produces a β-lactamase sensitive to clavulanic acid may be used as indicator organism.)

Zones of inhibition which appeared around the S. lividans colonies upon incubation overnight at 30°C were evidence of clavulanic acid production. Clavulanic acid-producing colonies were found amongst these initial S. lividans transformants at a frequency of about 12%. When plasmid DNA was isolated from one of these clavulanic acid-producing transformants and re-introduced into S. lividans, the frequency of clavulanic acid production in these 2nd round transformants was about 40-45%. Figure 6 shows a photograph of an agar plate bearing 2nd. round transformants. Zones of inhibition are seen as clear areas in the agar; these appear on the photograph as dark circular areas.

Example 5Sequencing of 15 kb DNA fragment

Ordered sets of deletions were generated as described in Example 1 using fragments of the DNA insert from the cosmid clone K6L2 (Figure 9) and subcloned into the E. coli plasmids pUC118 and pUC119. Overlapping fragments were chosen which extended from the end of the pcbC gene downstream for a distance of about 15 kb ending at the BglII site. The deletion generated fragments were sequenced in both orientations as described in Example 1. The sequence is shown in Figure 2.

The present invention is not limited to the features of the embodiments described herein, but includes all variations and modifications within the scope of the claims.

TABLE 1

ORF #	Start location (bp)	End location (bp)	Length (bp)	Size of ORF (aa residues)
1*	109	1764	1656	552
2	2216	3937	1722	574
3	3940	5481	1542	514
4	5654	6595	942	314
5	6611	7588	978	326
6	7895	9076	1182	394
7	9241	10 908	1668	556
8*	10 998	12 296	1299	433
9*	12 622	13 365	744	248
10	13 769	14 995	1227	409

* Asterisks denote ORFs which are oriented in the opposite direction.

The embodiments of the invention in which an exclusive property or privilege is claimed are defined as follows:

- 5 1. An isolated genomic DNA molecule comprising the nucleotide sequence of Figure 2 (Sequence ID No.:1).
2. An isolated DNA molecule having the nucleotide sequence of nucleotides 2033 to 13636 of Figure 2
10 (Sequence ID No.:20).
3. An isolated DNA molecule having the nucleotide sequence of nucleotides 109 to 1764 of Figure 2 (Sequence ID No.:21).
- 15 4. An isolated DNA molecule having the nucleotide sequence of nucleotides 2216 to 3937 of Figure 2 (Sequence ID No.:22).
- 20 5. An isolated DNA molecule having the nucleotide sequence of nucleotides 3940 to 5481 of Figure 2 (Sequence ID No.:23).
6. An isolated DNA molecule having the nucleotide sequence of nucleotides 5654 to 6595 of Figure 2
25 (Sequence ID No.:24).
7. An isolated DNA molecule having the nucleotide sequence of nucleotides 6611 to 7588 of Figure 2
30 (Sequence ID No.:25).
8. An isolated DNA molecule having the nucleotide sequence of nucleotides 7895 to 9076 of Figure 2
(Sequence ID No.:26).

9. An isolated DNA molecule having the nucleotide sequence of nucleotides 9241 to 10908 of Figure 2 (Sequence ID No.:27).
- 5 10. An isolated DNA molecule having the nucleotide sequence of nucleotides 10998 to 12296 of Figure 2 (Sequence ID No.:28).
11. An isolated DNA molecule having the nucleotide sequence of nucleotides 12622 to 13365 of Figure 2 (Sequence ID No.:29).
- 10 12. An isolated DNA molecule having the nucleotide sequence of nucleotides 13769 to 14995 of Figure 2 (Sequence ID No.:30).
- 15 13. An isolated DNA molecule comprising a nucleotide sequence encoding the amino acid sequence of Figure 10.
- 20 14. An isolated DNA molecule comprising a nucleotide sequence encoding the amino acid sequence of Figure 11.
- 25 15. An isolated DNA molecule comprising a nucleotide sequence encoding the amino acid sequence of Figure 12.
16. An isolated DNA molecule comprising a nucleotide sequence encoding the amino acid sequence of Figure 13.
- 30 17. An isolated DNA molecule comprising a nucleotide sequence encoding the amino acid sequence of Figure 14.
- 35

18. An isolated DNA molecule comprising a nucleotide sequence encoding the amino acid sequence of Figure 15.
- 5 19. An isolated DNA molecule comprising a nucleotide sequence encoding the amino acid sequence of Figure 16.
- 10 20. An isolated DNA molecule comprising a nucleotide sequence encoding the amino acid sequence of Figure 17.
- 15 21. An isolated DNA molecule comprising a nucleotide sequence encoding the amino acid sequence of Figure 18.
- 20 22. An isolated DNA molecule comprising a nucleotide sequence encoding the amino acid sequence of Figure 19.
23. An isolated protein having the amino acid sequence of Figure 10.
24. An isolated protein having the amino acid sequence of Figure 11.
- 25 25. An isolated protein having the amino acid sequence of Figure 12.
- 30 26. An isolated protein having the amino acid sequence of Figure 13.
27. An isolated protein having the amino acid sequence of Figure 14.
- 35 28. An isolated protein having the amino acid sequence of Figure 15.

29. An isolated protein having the amino acid sequence of Figure 16.
30. An isolated protein having the amino acid sequence of Figure 17.
31. An isolated protein having the amino acid sequence of Figure 18.
32. An isolated protein having the amino acid sequence of Figure 19.
33. A recombinant vector comprising a DNA molecule in accordance with any of claims 1 to 22.
34. A host transformed with a recombinant vector comprising a DNA molecule in accordance with any of claims 1 to 22.
35. A host transformed with a recombinant vector in accordance with claim 2 wherein the host is a Streptomyces.
36. A host in accordance with claim 35 which is S. lividans.
37. A process for producing clavulanic acid in a non-clavulanate-producing host comprising transforming the host with a DNA molecule in accordance with claim 2 and culturing the host under suitable conditions to produce clavulanic acid.
38. A process for producing clavulanic acid in accordance with claim 37 wherein the host is S. lividans.
39. A process for enhancing clavulanic acid production in a clavulanate-producing host comprising

transforming the host with a DNA molecule comprising a nucleotide sequence encoding one or more of the enzymes of the clavulanate synthetic pathway.

2108113

N-terminal amino acid sequence of CLA	Met	Glu	Arg	Ile	Asp	Ser	His	Val	Ser	Pro	Arg	His	Asp	(Asp)
	Tyr	Ala	Gln	Ile	Pro	Thr	Phe	Met	Arg	(Leu)	Pro	His	Asp	(Asp)
Potential codons (DNA)	TAT	GCT	CAA	ATT	CCT	ACT	TTT	ATG						
	C	C	G	C	C	C	C	C						
		A		A	A	A								
		G			G	G								
Probe made = 24-mer oligonucleotide with 8-fold degeneracy	TAC	GCC	CAG	ATC	CCC	ACC	TTC	ATG						
		G			G	G								
Actual DNA sequence	TAC	GCA	CAG	ATC	CCC	ACC	TTC	ATG						

FIGURE 1

Simon, M. B. 1977

2108113

FIGURE 2 - 1

	1	10	20	30	40	50	60	
1	gcggaaccgg	ccgccccctga	gcggggcggc	cgggaaggaa	acgggcccgt	cgccccctcg	60	
61	ggagggggcg	gccggccccgt	ccggtgcgcg	cggtggggtgc	ggcgcgggTC	AGCCGGCCGC	120	End of ORF 1
121	GAGGTTGCTG	AGGAACTTCG	CGGCGACGGG	GCCCCGCTCG	GCGCCGCCCG	ACCCGCCGTC	180	
181	CTCCAGCAGG	ACCGACCAGG	CGATGTTCCG	GTCGCCCTGG	TAGCCGATCA	TCCAGGCGTG	240	
241	CGTCTTCGGC	GGCTTCTCGG	TGCCGAACTC	GGCGGTACCG	GTCTTGCGGT	GCGGCTGTCC	300	
301	GCCGAGGCCC	CGCAGGGCGT	CGCCGGCGCC	GTCGGTGACG	GTCGAACGCA	TCATGGAACG	360	
361	CAGCGAGTCG	ACGATGCCCG	GGGCCATCCG	GGGGCCCTGG	TGCGGCTTCT	TGACCGCGTC	420	
421	GGGCACCAGC	ACGGGCTGCT	TGAACTCGCC	CTGCTTGACG	GTGGCGGCGA	TGGAGGCCAT	480	
481	CACCAGGGGC	GACGCCTCGA	CCCTGGCCTG	TCCGATGGTG	GACGCGGCCT	TGTCGTTCTC	540	
541	GCTGTTGGAG	ACGGGGACGC	TGCCGTGCAA	GGTGGAGGCG	CCGACGTCCC	AGGTGCCGCC	600	
601	GATGCCGAAG	GCTTCGGCGG	CCTGCTTCAG	GCTGGACTCG	GAGAGCTTGC	TGCGGGAGTT	660	
661	GACGAAGAAC	GTGTTGCAGG	AGTGGGCGAA	GCTGTCCCGG	AAGGTCGAGC	CCGCGGGCAG	720	
721	CGTGAAGTGG	TCCTGGTTCT	CGAAGCTCTG	GCCGTTGACA	TGGGCGAACT	TCGGGCAGTC	780	
781	GGCCCCGTCC	TCCGGGTTCA	TCCCCTGCTG	GAGCAGGGCC	GCGGTGGTGA	CCACCTTGAA	840	
841	GGTGGAGCCG	GGCGGGTAGC	GGCCCTCCAG	CGCGCGGTTT	ATGCCGGAGG	GCACGTTGCG	900	
901	GGCGGCCAGG	ATGTTGCCGG	TGGCGGGGTC	GACGGCGACG	ATCGCCGCGT	TCTTCTTCGA	960	
961	GCCCTCCAGG	GCCGCCGCGG	CGGCGGACTG	GACCCGCGGG	TCGATGGTGG	TCTTCACCGG	1020	
1021	CTTGCCCTCG	GTGTCCTTGA	GGCCGGTGAG	CTTCTTGACC	ACCTGGCCGG	ACTCACGGTC	1080	
1081	CAGGATCACG	ACCGAGCGCG	CCGCGCCGGA	GCCGCCGGTG	AGCTGCTTGT	CGTAGCGGGA	1140	
1141	CTGGAGGCCC	GCCGAGCCCT	TGCCGGTCCT	GGGGTCGACC	GCGCCGATGA	TGGAGGCGGC	1200	
1201	CTGGAGGACA	TTGCCGTTGG	CGTCGAGGAT	GTCCGCGCGC	TCCCAGGACT	TGAGGGCGAG	1260	
1261	GGTCTGCCCC	GGAACCATCT	GCGGATGGAT	CATCTCGGTG	TTGAACGCGA	CCTTCCACTC	1320	
1321	CTTGCCGCGG	CCGACGACCT	TCGCGGTGGA	GTCCAGGCGG	TACTCCCCGG	CCCCGGGGAG	1380	
1381	GGTCATTCTG	ACGGTGAACG	GATCTCCAC	CTCGCCCTCG	GGGTTCTTCT	CCCCGGTCTT	1440	
1441	GGCGGTGATC	TCCGTCTTCG	TCGGCTTGAG	GTTGGTCATG	ACGGATTTGA	TCAGCGACTC	1500	
1501	GGCGTTGTCC	GGGGTGTCCG	TCAGCCCGGC	GGCGTCGGG	GCGTCGCCCT	TCTCCAGGC	1560	

Sim; M. Bunnif

FIGURE 2 - 2

1561 GCCGAGGAAG GTGTGAACT GTCCGGCCGC CGCTCCACC TCGGGGTCGC CCGAATCCTT 1620
 1621 CTCGTCGGCA ACCAGGCTGG TGTAACCCCA ATAGCCGAGC CCCACCGTCA CGGCCAGCCC 1680
 1681 GGCGACCACC GCGGTGGCCG CCCGGCCACG GGAGCGGCGC CTGCCCTGCG GCGGGTCATC 1740
 1741 GCCATAGTTG TCGGAATGCG TCATggggcc aggcctatgcg ggccgacctct ttccctcctc 1800
 1801 cccggatacc gcgtttcagg acagtcaagg ggccgaacgg agggctggac cagccgctca 1860
 1861 ggggcccgtt cccacccctt ggggggaagc ggcacccgga aggtgaccga ggcaacatcc 1920
 1921 atggaaaggg gagcgaatcg gtcgccgagt tcaccgcgat tggagtagac ctctgaaagc 1980
 1981 gtgacagcgg ggagtagcga caaacgcgtc agacccctga agggaattga ctgaattcga 2040
 2041 gtcctcgggt tcggcgacgg atgggcggtt cggccacgca ccgtcactct tcgtccctc 2100
 2101 ttcaaaagaa ctcccgatac gtggagaaga gagcgtgaag agcgcgtccg gtcagggttg 2160
 2161 ccgagaaccg tccaccatga cggagcctgg tactgacgga gtcgggagac cgctcATGTC 2220
 2221 CCGTGTATCG ACCGCCCCCA GCGGCAAGCC TACCGCCGCT CACGCCCTCC TGTCACGGTT 2280
 2281 GCCTGATCAC GGTGTGGGGA AGGTGTTTGG GGTGTCTGGC CGAGAGGCCG CGTCGATTCT 2340
 2341 CTTGACGAG GTCGACCCCA TCGACTTCGT TCTGACCCGC CACGAGTTCA CCGCGGGTGT 2400
 2401 CGCCGCTGAT GTCCTCGCGC GGATCACCGG TCGCCCCCAG GCGTGCTGGG CCACCCTGGG 2460
 2461 CCCCAGTATG ACCAACCTCT CCACCGGTAT CGCCACGTCC GTCCTGGACC GCTCGCCGGT 2520
 2521 CATCGCGCTC GCCGCGCAGT CGGAGTCGCA CGACATCTTC CCGAACGACA CCCACCAGTG 2580
 2581 CCTGGACTCG GTGGCGATCG TCGCCCCGAT GTCCTTGAC GCCGTGGAGC TCCAGCGGCC 2640
 2641 CCACGAGATC ACCGACCTCG TCGACTCCGC CGTGAACGCG GCCATGACCG AGCCGGTCGG 2700
 2701 GCCCTCCTTC ATCTCCCTCC CCGTGGACCT GCTCGGCTCC TCCGAGGGCA TCGACACCAC 2760
 2761 CGTCCCCAAC CCGCCGGCGA ACACCCCGGC GAAACCGGTC GGCGTCGTCG CCGACGGCTG 2820
 2821 GCAGAAGGCC GCCGACCAGG CCGCCGCCCT GCTCGCCGAG GCCAAGCACC CCGTGCTCGT 2880
 2881 CGTCGGAGCG GCCGCGATCC GCTCGGGCGC CGTCCCGCG ATCCGCGCCC TGGCCGAGCG 2940
 2941 CCTGAACATC CCGGTCATCA CGACCTACAT CGCCAAGGGT GTCCTGCCGG TCGGCCACGA 3000
 3001 GCTGAACTAC GGCGCCGTCA CCGGCTACAT GGACGGCATC CTCAACTTCC CCGCGCTCCA 3060
 3061 GACCATGTTT GCCCCGGTGG ACCTCGTCCT CACCGTCGGC TACGACTACG CCGAGGACCT 3120
 3121 GCGCCCGTCC ATGTGGCAGA AGGGCATCGA GAAGAAGACC GTCCGTATCT CCCCAGCGGT 3180
 3181 CAACCCGATC CCCCAGGTCT ACCGGCCCGA CGTCGACGTC GTCACCGACG TCCTCGCCTT 3240

Sim; M. Barmy

FIGURE 2 - 3

3241 CGTGGAGCAC TTCGAGACCG CGACCGCCTC CTTCCGGGGCC AAGCAGCGCC ACGACATCGA 3300
 3301 GCCGCTGCGC GCCCGGATCG CGGAGTTCCT GGCCGACCCG GAGACCTACG AGGACGGCAT 3360
 3361 GCGCGTCCAC CAGGTTCATCG ACTCCATGAA CACCGTCATG GAGGAGGCCG CCGAGCCCGG 3420
 3421 CGAGGGCACG ATCGTCTCCG ACATCGGCTT CTTCCGTAC TACGGTGTGC TCTTCGCCCC 3480
 3481 CGCCGACCAG CCCTTCGGCT TCCTCACCTC GGCGGGCTGC TCCAGCTTCG GCTACGGCAT 3540
 3541 CCCC GCCGCC ATCGGGCCCC AGATGGCCCC CCCGGACCAG CCGACCTTCC TCATCGCGGG 3600
 3601 TGACGGCGGC TTCCACTCCA ACAGCTCCGA CCTGGAGACC ATCGCCCGGC TCAACCTGCC 3660
 3661 GATCGTGACC GTCGTCTGCA ACAACGACAC CAACGGCCTG ATCGAGCTGT ACCAGAACAT 3720
 3721 CGGTCAACCAC CGCAGCCACG ACCCGGGCGT CAAGTTCGGC GGCCTCGACT TCGTCGCGCT 3780
 3781 CGCCGAGGCC AACGGTGTGC ACGCCACCCG CGCCACCAAC CGCGAGGAGC TGCTCGCGGC 3840
 3841 CCGTGC CAAG GGTGCCGAGC TGGGTCTGTC GTTCTCATC GAGGTCCCGG TCAACTACGA 3900
 3901 CTTCCAGCCG GGCGGCTTCG GCGCCCTGAG CATCTGATcA TGGGGGCACC GGTTCCTCCG 3960
 3961 GCTGCCTTCG GGTTCTGGC CTCCGCCCGA ACGGGCGGGG GCCGGGCCCC CGGCCCGGTC 4020
 4021 TTCGCGACCC GGGGCAGCCA CACCGACATC GACACGCCCC AGGGGGAGCG CTCGCTCGCG 4080
 4081 GCGACCCTGG TGCACGCCCC CTCGGTCGCG CCCGACCGCG CGGTGGCGCG CTCCTCACC 4140
 4141 GGCGCGCCCA CCACCGCGGT GCTCGCCGGT GAGATCTACA ACCGGGACGA ACTCCTCTCC 4200
 4201 GTGCTGCCCC CCGGACCGCG GCCGGAGGGG GACGCGGAGC TGGTCTGCG GCTGCTGGAA 4260
 4261 CGCTATGACC TGCATGCCTT CCGGTGGTG AACGGGCGCT TCGCGACCGT GGTGCGGACC 4320
 4321 GGGGACCGGG TCCTGCTCGC CACCGACCAC GCGGGTTCGG TGCCGCTGTA CACCTGTGTG 4380
 4381 GCGCCGGGCG AGGTCCGGGC GTCCACCGAG GCCAAGGCGC TCGCCGCGCA CCGCGACCCG 4440
 4441 AAGGGCTTCC CGCTCGCGGA CGCCCGCCGG GTCGCCGGTC TGACCGGTGT CTACCAGGTG 4500
 4501 CCCGCGGGCG CCGTGATGGA CATCGACCTC GGCTCGGGCA CCGCCGTAC CCACCGCACC 4560
 4561 TGGACCCCGG GCCTCTCCCG CCGCATCCTG CCGGAGGGCG AGGCCGTGCG GGCCGTGCGG 4620
 4621 GCCGCGCTGG AGAAGGCCGT CGCCAGCGG GTCACCCCG GCGACACCCC GTTGGTGGTG 4680
 4681 CTCTCCGGCG GAATCGACTC CTCCGGGGTC GCGGCCTGTG CGCACCGGGC GGCCGGGGAA 4740
 4741 CTGGACACGG TGTCCATGGG CACCGACACG TCCAACGAGT TCCGCGAGGC CCGGGCGGTC 4800
 4801 GTCGACCATC TGCGACCCCG GCACCGGGAG ATCACCATCC CGACCACCGA GCTGCTGGCG 4860

Sim; M. Bunnif

FIGURE 2 - 4

4861 CAGCTCCCGT ACGCGGTGTG GGCCTCCGAG TCGGTGGACC CGGACATCAT CGAGTACCTG 4920
 4921 CTCCCCCTGA CAGCGCTCTA CCGGGCGCTC GACGGGCCGG AGCGCCGCAT CCTCACC GGG 4980
 4981 TACGGCGCGG ACATCCCCCT CGGGGGCATG CACCGCGAGG ACCGGCTGCC CGCGCTGGAC 5040
 5041 ACCGTTCTCG CGCACGACAT GGCCACCTTC GACGGGCTGA ACGAGATGTC CCCGGTGCTG 5100
 5101 TCCACGCTGG CGGGGCACTG GACCACCCAC CCGTACTGGG ACCGGGAGGT CCTCGATCTG 5160
 5161 CTGGTCTCGC TGGAGGCCGG GCTCAAGCGG CGGCACGGCC GGGACAAGTG GGTGCTGCGC 5220
 5221 GCCGCGATGG CCGACGCCCT CCCGGCGGAG ACCGTCAACC GGCCCAAGCT GGGCGTCCAC 5280
 5281 GAGGGCTCGG GCACCACGTC CTCGTTCTCC CGGCTGCTGC TGGACCACGG TGTCGCCGAG 5340
 5341 GACCGCGTCC ACGAGGCGAA GCGGCAGGTG GTGCGCGAGC TGTTGATCT CACGGTCGGG 5400
 5401 GGCGGACGGC ACCCCTCCGA GGTGGACACC GACGATGTGG TGCCTCCGT GGCCGACCGG 5460
 5461 ACCGCGCGGG ^{End of ORF 3--}GGGCGGCCTA Gtccccccac ggggagcccc cgggacgccc gacccgcgcy 5520
 5521 ggacccgtac ccggggccgc ccgcggactc cggcgacccg gcacccctgt cccccacccg 5580
 5581 ttgacgaccg tggccctcg gccctcgagg cccctgacga ccgtcgcccc attcccagga 5640
 5641 gggagctgaa ^{Beginning of ORF 4--}agcGTGGAGC GCATCGACTC GCACGTTTCA CCCCCTACG CACAGATCCC 5700
 5701 CACCTTCATG CGCCTGCCGC ACGATCCCCA GCCCCGCGGC TATGACGTGG TGGTCATCGG 5760
 5761 AGCCCCCTAC GACGGGGGCA CCAGCTACCG TCCCGGCGCC CGGTTGCGCC CCCAGGCCAT 5820
 5821 CCGCAGTGAG TCGGGCCTCA TCCACGTTGT CGGCATCGAC CGGGGCCCCG GCACGTTTCA 5880
 5881 CCTGATCAAC TGTGTCGACG CCGGGGACAT CAATCTGACG CCGTTCGACA TGAACATCGC 5940
 5941 GATCGACACG GCGCAGAGCC ATCTGTCGGG CCTGCTGAAG GCCAACGCCC CCTTTCTGAT 6000
 6001 GATCGGCGGC GACCACTCGC TGACGGTGGC CGCCTGCGC GCGGTGCGCG AGCAGCACGG 6060
 6061 CCCGCTCGCC GTGGTGACC TGGACGCGCA CTCCGACACC AACC CGGCCT TCTACGGGGG 6120
 6121 CCGGTACCAC CACGGCACCC CTTCCGGCA CGGGATCGAC GAGAAGCTGA TCGACCCGGC 6180
 6181 GGCGATGGTC CAGATCGGCA TCCGGGGCCA CAACCCGAAG CCGGACTCGC TCGACTACGC 6240
 6241 CCGGGGCCAC GCGTCCGGG TGGTCACGGC GGACGAGTTC GGCGAGCTGG GGGTGGGCGG 6300
 6301 GACCGCCGAC CTCATCCGCG AGAAGGTGCG CCAGCGGCCC GTGTACGTCT CGGTGACAT 6360
 6361 CGACGTGGTC GACCCCGCCT TCGCCCCCGG TACGGGCACG CCCGCGCCGG GCGGGCTCCT 6420
 6421 CTCGCGCGAG GTGCTGGCGC TGCTGCGCTG CGTGGGTGAC CTGAAGCCGG TCGGCTTCGA 6480
 6481 CGTGATGGAG GTGTCACCCC TCTACGACCA CGGCGGGATC ACTTCGATCC TGGCCACGGA 6540

Simon, M. Barry

2108113

FIGURE 2 - 5

6541 GATCGGTGCG GAACTGCTCT ACCAGTACGC CCGAGCCCAC ^{End of ORF 4-->} AGAACCCAGT TGTGAaggag 6600
 6601 acatcgtgtc ^{Beginning of ORF 5-->} ATGGCCTCTC CGATAGTTGA CTGCACCCCG TACCGCGACG AGCTGCTCGC 6660
 6661 GCTCGCCTCC GAGCTTCCCG AGGTGCCGCG CGCGGACCTC CATGGCTTCC TCGACGAGGC 6720
 6721 GAAGACGCTG GCCGCCCCGTC TCCCGGAGGG GCTGGCCGCC GCTCTCGACA CCTTCAACGC 6780
 6781 CGTGGGCAGC GAGGACGGTT ATCTGCTGCT GCGCGGGCTG CCCGTCGACG ACAGCGAGCT 6840
 6841 GCCCGAGACG CCGACCTCCA CCCC GGCCCC GCTGGACCGC AAGCGGCTGG TGATGGAGGC 6900
 6901 CATGCTCGCG CTGGCCGGCC GCCGGCTCGG TCTGCACACG GGGTACCAGG AGCTGCGCTC 6960
 6961 GGGCACGGTC TACCACGACG TGTACCCGTC GCCCGGCGCG CACTACCTGT CCTCGGAGAC 7020
 7021 CTCCGAGACG CTGCTGGAGT TCCACACGGA GATGGCGTAC CACATCCTCC AGCCGAACCTA 7080
 7081 CGTCATGCTG GCCTGCTCCC GCGCGGACCA CGAGAACCGG GCGGAGACGC TGGTCGGCTC 7140
 7141 GGTCCGCAAG GCGCTGCCCC TGCTGGACGA GAAGACCCGG GCCCGTCTCT TCGACCGCAA 7200
 7201 GGTGCCCTGC TGCGTGGACG TGGCCTTCCG CGGCGGGGTC GACGACCCGG GCGCGATCGC 7260
 7261 CAACGTCAAG CCGCTCTACG GGGACGCGAA CGACCCGTTT CTCGGGTACG ACCGCGAGCT 7320
 7321 GCTGGCGCCG GAGGACCCCG CGGACAAGGA GGCCGTCGCC CATCTGTCCC AGGCGCTCGA 7380
 7381 CGATGTGACC GTCGGGGTGA AGCTCGTCCC CGGTGACGTC CTCATCATCG ACAACTTCCG 7440
 7441 CACCACGCAC GCGCGGACGC CGTTCTCGCC CCGCTGGGAC GGGAAGGACC GCTGGCTGCA 7500
 7501 CCGCGTCTAC ATCCGCACCG ACCGCAATGG ACAGCTCTCC GGCGGCGAGC GCGCGGGCGA 7560
 7561 CACCATCTCG ^{End of ORF 5-->} TTCTCGCCGC GCCGCTGAGc ccggtcctcc gagggccctgg gccccggcgc 7620
 7621 cgggaaccggc tcccggctct gccccctcac ccgcccgcgc ggtgaggggg caggccccctt 7680
 7681 tgtgccgggt gccgtgcgtc ctgcgaggggt gccggggcgc ggggggacggc ggagggtgcc 7740
 7741 ggcgccgggg tgccgtgcgc cggccgtggg tgctgtacag cactccgtgt gccgtgcgcc 7800
 7801 accccgtgca taaatttgcc actctatggg aaataatgca gagtgcgacg ggtgaggccg 7860
 7861 tcgccgtgcc ctttccgtga caggagacgc ^{Beginning of ORF 6-->} tgacATGTCC GACAGCACAC CGAAGACGCC 7920
 7921 CCGGGGATTG GTGGTGCACA CGGCGCCGGT GGGCCTGGCC GACGACGGCC GCCACGACTT 7980
 7981 CACCGTCCTC GCCTCCACCG CCCC GGCCAC CGTGAGCGCC GTCTTCACCC GCTCCCCTT 8040
 8041 CGCCGGGCCG AGCGTCGTGC TGTGCCGGGA GGCGGTGGCC GACGGGCAGG CGCGCGGTGT 8100
 8101 GGTGGTGCTG GCCCGCAACG CGAATGTCGC GACCGGCCTG GAGGGCGAGG AGAACGCGCG 8160

Simon, M. Barry

FIGURE 2 - 6

8161 CGAGGTGCGC GAGGCCGTCG CCCGGGCCCT CGGGTGCCG GAGGGCGAGA TGCTGATCGC 8220
 8221 CTCCACCGGG GTGATCGGCC GGCAGTACCC GATGGAGAGC ATCCGGGAGC ACCTCAAGAC 8280
 8281 GCTGGAGTGG CCCGCCGGGG AGGGCGGCTT CGACCGCGCG GCCCGCGCCA TCATGACGAC 8340
 8341 CGACACCCGG CCCAAGGAGG TCCGGGTGAG CGTCGGCGGG GCGACCCTCG TGGGCATCGC 8400
 8401 CAAGGGCGTC GGCATGCTGG AGCCCGACAT GGGCAGCTG CTGACCTTCT TCGCCACGGA 8460
 8461 CGCCCGGCTG GACCCGGCCG AGCAGGACCG CCTCTTCCGC CGGGTCATGG ACCGCACCTT 8520
 8521 CAACGCGGTC AGCATCGACA CCGACACCTC CACCAGCGAC ACGGCGGTGC TGTTGCCAA 8580
 8581 CGGCCTGGCG GCGGAGGTG ACGCCGGGGA GTTCGAGGAG GCGCTGCACA CGGCGGCGCT 8640
 8641 GGCCTGGTC AAGGACATCG CGAGCGACGG CGAGGGCGCG GCCAAGCTGA TCGAGGTCCA 8700
 8701 GGTCACCGGC GCCCGCGACG ACGCCAGGC CAAGCGGGTC GGCAAGACCG TCGTCAACTC 8760
 8761 CCCGTTGGTG AAGACCGCCG TGCACGGCTG CGACCCCAAC TGGGGCCGGG TCGCCATGGC 8820
 8821 GATCGGCAAG TGCTCGGACG ACACCGACAT CGACCAGGAG CGGGTGACGA TCCGCTTCGG 8880
 8881 CGAGGTCGAG GTCTATCCGC CGAAGGCCCG GGGCGACCAG GCCGACGACG CGCTGCGGGC 8940
 8941 CGCCGTCGCG GAGCATCTGC GGGGCGACGA GGTGGTCATC GGGATCGACC TCGCCATCGC 9000
 9001 GGACGGGGCC TTCACGTCT ACGGCTGCGA CCTCACCAGG GGCTATGTCC GGCTGAACTC 9060
 9061 GGAGTACACC ACCTGATccc cggacagggg acggggccgc gccccgttcc ctgtccgctc 9120
 9121 ccgtcccggtg tggttatacc gaccgttccc cggctatgcg cacgggacgg agcggccccc 9180
 9181 gcccgggccc gcccggcgcc acgatgaggg gcgatgcaag gtgacgaggg caggagggac 9240
 9241 ATGGAGACCA CTCGGTCGAC GACCGCGGAC GAGGGCTTCG ACGCCGGGGT ACGGGGAGTG 9300
 9301 GTCGCGCCGA CCGACGCCCC GGGCGGGACG CTGCGGCTGG TCCGCACGGA CGACTTCGAC 9360
 9361 TCGCTCGACC CCGGCAACAC GTACTACGCC TACACCTGGA ACTTCCTCCG GTCATCGGC 9420
 9421 CGGACGCTGG TCACCTTGA CACCGCGCCG GGCAAGGCGG GCCAGCGGCT CGTGCCCGAC 9480
 9481 CTCGCCGAGT CGCTGGGCGA GTCCTCCGAG GACGGCCGGG TCTGGACCTA CCGGCTGCGC 9540
 9541 GAGGGCCTGC GCTACGAGGA CGGCACGCCG GTCGTCTCGG CCGACATCAA GCACGCCATC 9600
 9601 GCCCGCAGCA ACTACGGCAC CGATGTCCTG GGGCGCGGTC CGACCTACTT CCGCCACCTC 9660
 9661 CTGGGCACCG AGTACGGCGG CCCCTGGCGG GAGCCGGACG CCGACGGACC GGTGACGCTG 9720
 9721 GAGACCCCGG ACGAGCGGAC GCTGGTCTTC CGGCTGCGGG AGCCGTTGCG GGGGATGGAT 9780
 9781 CTGCTGGCGA CCATGCCGTC CACCACCCCC GTGCCGCGCG ACCGGGACAC CGGCGCCGAG 9840

Simon; M. B.

FIGURE 2 - 7

9841 TACCGGCTGC GGCCCGTGGC GACCGGCCCG TACCGGATCG TCTCGTACAC CCGGGGCGAG 9900
 9901 CTGGCCGTCC TGGAGCCCAA TCCGCACTGG GACCCCGAGA CCGACCCGGT GCGCGTCCAG 9960
 9961 CGCGCCTCCC GGATCGAGGT GCACCTCGGC AAGGACCCGC ACGAGGTGGA CCGCATGCTG 10020
 10021 CTGGCGGGCG AGGCCCATGT GGACCTCGCG GGCTTCGGTG TGCAGCCCGC GGCCAGGAG 10080
 10081 CGCATCCTCG CCGAGCCGGA GCTGCGCGCG CACGCGGACA ACCCGCTGAC CGGCTTCACC 10140
 10141 TGGATCTACT GCCTGTCGAG CCGGATCGCC CCGTTCGACA ATGTGCACTG CCGGCGGGCC 10200
 10201 GTGCAGTTCG CCACCGACAA AGCGGCCATG CAGGAGGCGT ACGGCGGGCG GGTGGGCGGC 10260
 10261 GACATCGCGA CCACCCTGCT GCCCCGACC CTCGACGGCT ACAAGCACTT CGACCGCTAC 10320
 10321 CCGGTCGGCC CCGAGGGCAC CGGCGACCTG GAGGCCGCCC GCGCCGAGCT GAAGCTGGCC 10380
 10381 GGGATGCCCC ACGGCTTCCG CACCAGGATC GCCGCCCGCA AGGACCGGCT CAAGGAGTAC 10440
 10441 CGGGCCGCCG AGGCGCTGGC CGCCGGGCTC GCCCGGGTCG GCATCGAGGC GGAGGTGCTG 10500
 10501 GACTTCCCGT CGGGCGACTA CTTGACCGC TACGGCGGCT GCCCGAGTA TCTGCGCGAG 10560
 10561 CACGGGATCG GGATCATCAT GTTCGGCTGG GGCGCCGACT TCCCCGACGG ATACGGCTTC 10620
 10621 CTCCAGCAGA TCACCGACGG GCGCGCGATC AAGGAGCGCG GCAACCAGAA CATGGGCGAG 10680
 10681 CTGGACGACC CGGAGATCAA CGCGCTGCTG GACGAGGGGG CGCAGTGCGC CGACCCGGCG 10740
 10741 CGGCGCGCGG AGATCTGGCA CCGCATCGAC CAGCTCACGA TGGACCACGC GGTTCATCGTT 10800
 10801 CCGTATCTGT ACCCGCGGTC CCTGCTCTAC CGGCACCCGG ACACCCGCAA CGCCTTCGTC 10860
 10861 ACCGGCTCCT TCGGGATGTA CGACTACGTG GCGCTCGGCG CGAAGTGAgc acgggggtccg 10920
 10921 gccccgggac cgtatgtccc ggggcccggac cccgcccgtt ccccgcccgg tccggtccgg 10980
 10981 acccggtcgc ggcccgcTCA GCCGGACATC CGGGCCCCGG CCGCGACCCC GCGCCGGATC 11040
 11041 GGCCAGTGGC CCTGCGCCAG GGGCCGTTCC ACGCTGCGGC AGGCGAGAGC GGCCTCGCGG 11100
 11101 AACTCCGCCT CGTACAGCGC GAGCTGGCGC AGGAACTGCC GGGTCGGGCC GGTTCAGGCTG 11160
 11161 GTCCCCCGCG GGCTGCGCAG CAGCAGCCGG GCGCCGAGGG ACTGCTCCAG CCGGTGAATC 11220
 11221 CGGCGGGTGA GCGCCGACTG GCTGATCGAC AGCACCGCCG CGGCCCGGTT GATGCTGCCG 11280
 11281 TGCCGGGCCA CGGCCTGGAG CAGATGGAGA TCGTCCACAT CCAGTTTGCG GCCCTCGGCC 11340
 11341 TGGCCGGGCA CGGAGCCCTG GTCGGGTCCC GCCCGAAGC GGCGGGCGTC CGCGCCGGTG 11400
 11401 CGCTCCGCGT ACCACTGCGC CCACCAGGGC TCGTCCAGCA GGTGCGGGTG GTGTTGGCG 11460
 11461 AAGCGCCGGA GCTGGACCTC GCGGATCAGC GCGGCCAGCC GTCCCGCCAG CGCCCGGGGC 11520

Sim; M. Barmy

FIGURE 2 - 8

11521 ACGATGGTGG GGTGACGAG CAGACTCGTG GTGCGGCGCG GCGCTCCGC CAGGGAGCGG 11580
 11581 CGCACCAGCG AGGGGTCCTG CACCGCCGGG TGGGTGGGCG AGCCGAGACC TATCGCGTCC 11640
 11641 CCGCGGCGCA GGATGCCCCG GGCAACCGAT GCCCCGTGA TGTGGAGCCG GGTGGGCGCG 11700
 11701 GTGAGCCCGG CCAGCTGGAA GACACGTGTC ACCAGGATCT CCGAGCCGGG TCCCGTCTCG 11760
 11761 GACACCCAGG TCTCGTCCCG CAGATCGGCG AGCGAGACCT CCCGCCGGGC GGCCAGCGGA 11820
 11821 TGGTCCCGGG GCAGGATCAC CCACAGCGGG TCGTCCAGCA CCTCACAGGT GCGCACGGAC 11880
 11881 CGCTCCAGGC TGTGCCGGGG GGA CTCCAGGTGT AGGCCGCGTC CACCTGGTAG 11940
 11941 CCCGCCAGTT GGGCGGCGAC CTGGTGC GCGCTCGTCC GGACCGACAG CAGCAGGTCC 12000
 12001 AGCGAGGCCG CCGCGTCCTC CACCACCTCG TCGAGCAGGG GTTCCGTGGA GACCAGCGAC 12060
 12061 AGCACCTCCG GGGCGTCCAC GGCCTCGGAG CCATGGCCGA AGATATGCGT CCGCGCGGCC 12120
 12121 AGGTCGACCT GGTGGAAGAA CCGCCGCCCG GCGACGAGGA TGCGGGAGCC CGCGGTGGTC 12180
 12181 AGCCGGGCCG TGTGGCGGCT GCGCAGGGTC AGCGGGAGGC CGACGATCCG GTCCAGCCGG 12240
 12241 TCGAGTCTGC GCTCCACGGT GCCGTGCCGG ACACCGTCC GCCGGGCCAC TTCCATgagg 12300
 12301 tctccgcagt gtccaccgc gtccagtaaa gacagatcgc atcggctgac accagcagac 12360
 12361 gtcggttctg acccgagaga caatgtcgggt tcccttttcc gtcaaggact gtaccgctga 12420
 12421 attgtccgaa gtggtctctg aattgcttcg gaatcgatcc taggcagcgc cgctcttcgg 12480
 12481 atttctctcg ccgggaagcg gaacgcgccc ggccggatgg cgggcgcgct ccgggcgccc 12540
 12541 tcccgggaac gggggacggg gcacggcacg gcccggcacc cggctccggc gcgcggcgtg 12600
 12601 gacctggctg gcggacgggt gTCAGACCTG GTCGGTGGGG CGTATGAAGA TCTCGTGGAC 12660
 12661 GGTCGCGTGG TCGGCGCGCG TCACGGCGTA GCGGACCGCC TCCGCGATGT CCTGGGCCTG 12720
 12721 GAGCTTGCGG ATCTGGCTGA TCCGCTGCTC GTACATCTCC TTGGTGGCGG TGTGGGTGAT 12780
 12781 GTGGCCGCGC AGCTCCGTGT CGGTGGTGCC CGGCTCGATG ACGACGACCC GCACCCCGCG 12840
 12841 CTCGGTGACC TCCTGGCGCA GCGTCTCGCT GAACGCGTTC ACACCGAACT TCGTGGCCTG 12900
 12901 GTAGACGGCC GCGTTGCGGA CGTTCACCCG GCCCGCGATC GAGGACATCT GCACCACGGT 12960
 12961 GCCCTTGCTG CGCAGCAGAT GGGGAAGGGC CGCCCGGGTC ATGTACATCA GGCCAGGAG 13020
 13021 ATTGGTGTG ATCATCCGGG TCCAGTCGGT GGTGTCGGCG TCCTCCACCG GGCCGAGCAG 13080
 13081 CATGATCCCG GCGTTGTTGA CGAGGATGTC GAGGCCGCC AGCGCCTCGA CGGTGGAGGC 13140
 13141 GACGGCGGCG TCCACCCCT GCCGGTCGGC GACGTCGAGT TCGAGGACAT GGACCTTCGC 13200

Simon, M. Baum

2108113

FIGURE 2 - 9

13201 CCCGGCGGCG GTCAGCTCGT CACCCAGGGC GCGCAGCTTC TCGACCCGGC GCGCGGCGAT 13260
 13261 GGCCACGGCG GCGCCCTCGG CGGCCAGGGC GCGGGCCGTG GCCTCGCCGA TGCCCGAGCT 13320
 13321 CGCGCCCGTG ATGAGCGCGA CTTTCCCCTG GAGTGGCGAT GGCATcattt cctccacatg 13380
 13381 gtgctgcat cgtggtgagc gtatgaagaa ggggtgagac ctgccgtgcc ggggcgggtt 13440
 13441 ccgtacgccg gaccgttgcc gtgggcacgg ccgaccgggt acggatggcc gcagttcccc 13500
 13501 ggggagttcc cggggaatgg tgaataccgc ggcgtctccc gatggtcttc ggaggacacc 13560
 13561 cggggattca ccgggaatca gcggccggag ttctccccgt ccacggcaga cgtatcagc 13620
 13621 gtgcattcc ccggtgaatt cccttcggtg gaccgggtta tgactgtttc cgccgggtta 13680
 13681 tgccgcgccg cccggcggac cggccaccgc cccgggggct gcggcagatt gggcgccacg 13740
 13741 acatggcgcg agcagcgatc ggcggtggAT GATGAACGAG GCAGCGCCTC AGTCCGACCA 13800
 13801 GGTGGCACCG GCGTATCCGA TGCACCGGGT CTGCCCCGGT GACCCGCCGC CGCAACTGGC 13860
 13861 CGGGCTGCGG TCCCAGAAGG CCGCGAGCCG GGTGACGCTG TGGGACGGCA GCCAGGTGTG 13920
 13921 GCTGGTGACC TCGCAGCCG GGGCCCGGGC CGTCTGGGC GACCGCCGCT TCACCGCGGT 13980
 13981 GACGAGCGCG CCCGGCTTCC CGATGCTGAC CCGCACCTCC CAACTGGTGC GCGCCAACCC 14040
 14041 GGAGTCGGCG TCGTTCATCC GCATGGACGA CCCGCAGCAC TCCCGGCTGC GCTCGATGCT 14100
 14101 CACCCGGGAC TTCCTGGCCC GCCGCGCCGA GGCCTGCGC CCCGCGGTGC GGGAGCTGCT 14160
 14161 GGACGAGATC CTGGGCGGGC TGGTGAAGGG GGAGCGGCCG GTCGACCTGG TCGCCGGACT 14220
 14221 GACGATCCCG GTGCCCTCGC GGGTCATCAC CCTGCTCTTC GCGCGCGGTG ACGACCGCCG 14280
 14281 GGAGTTCATC GAGGACCGCA GCGCGGTCCT CATCGACCGC GGCTACACCC CGGAGCAGGT 14340
 14341 CGCCAAGGCC CGGGACGAAC TCGACGGCTA TCTGCGGGAG CTGGTCGAGG AGCGGATCGA 14400
 14401 GAACCCGGGC ACCGACCTGA TCAGCCGGCT CGTCATCGAC CAGGTGCGGC CGGGGCATCT 14460
 14461 GCGGGTCGAG GAGATGGTCC CGATGTGCCG GCTGCTGCTG GTGGCCGGTC ACGGCACCAC 14520
 14521 CACCAGCCAG GCGAGCCTGA GCCTGCTCAG CCTGCTCACC GACCCGGAGC TGGCCGGGCG 14580
 14581 CCTCACCAG GACCCGGCCC TGCTGCCCAA GGCGGTGAG GAGCTGCTGC GCTTCCACTC 14640
 14641 CATCGTGACG AACGGGCTGG CCCGTGCCGC GGTGGAGGAC GTCCAGCTCG ACGATGTGCT 14700
 14701 CATCCGGGCG GCGGAGGGCG TGGTGTGTC GCTGTGCGCG GGCAACCGGG ACGAGACGGT 14760
 14761 CTTCCCCGAC CCGGACCGGG TGGACGTGGA CCGCGACGCC CGCCGCCATC TCGCTTCGG 14820
 14821 CCACGGCATG CACCAGTGCC TGGGCCAGTG GCTGGCCCCG GTGGAGCTGG AGGAGATCCT 14880

Simon M. Lee

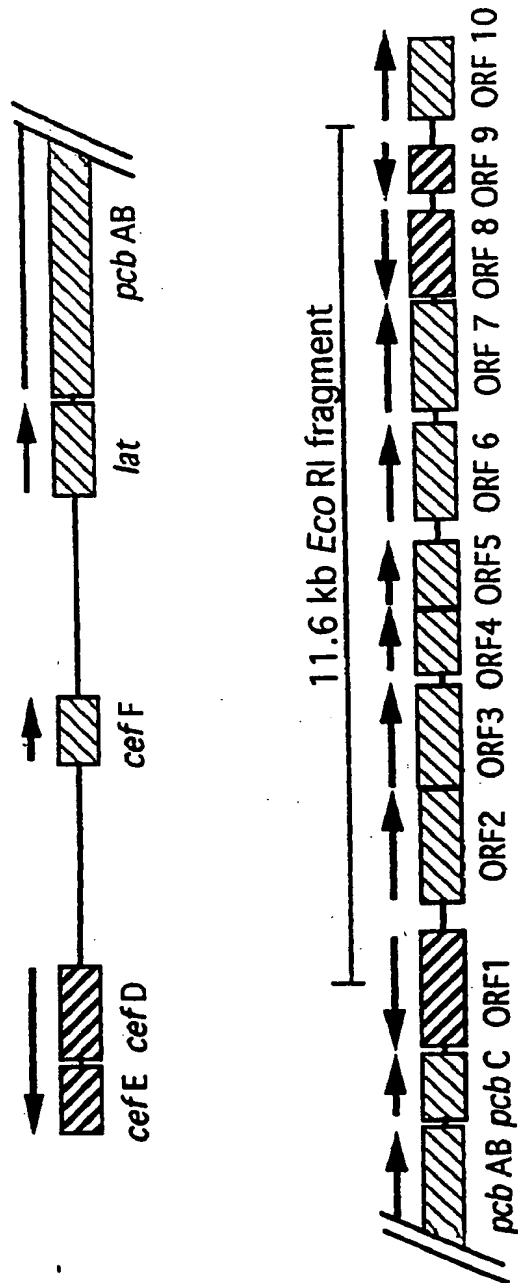
2108113

FIGURE 2 - 10

14881 CGCCGCGGTG CTGCGCTGGA TGCCCGGTGC CCGGCTCGCG GTGCCCTTCG AGGAGCTGGA 14940
14941 CTTCCGTCAT GAGGTGTCCA GTTACGGCCT CGGCGCCCTC CCGGTGACCT GGTGAgcggc 15000
15001 gtggagcggc tgaccgtcgt cctcgacgcg tcggcctgct gcgcgatggg gcgctgcgcg 15060
15061 gccacggccc ccgagatct 15079

| 10 | 20 | 30 | 40 | 50 | 60

Sim; M. Baum



ORF 4 = *c/a*

FIGURE 3

Simon, M. Barney

2108113

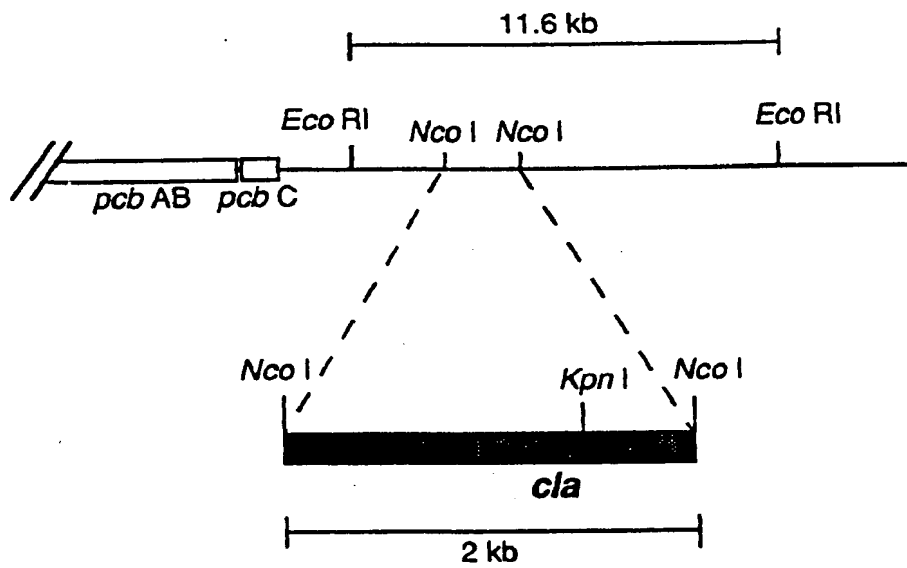


FIGURE 4

Sim; M. Baum

2108113

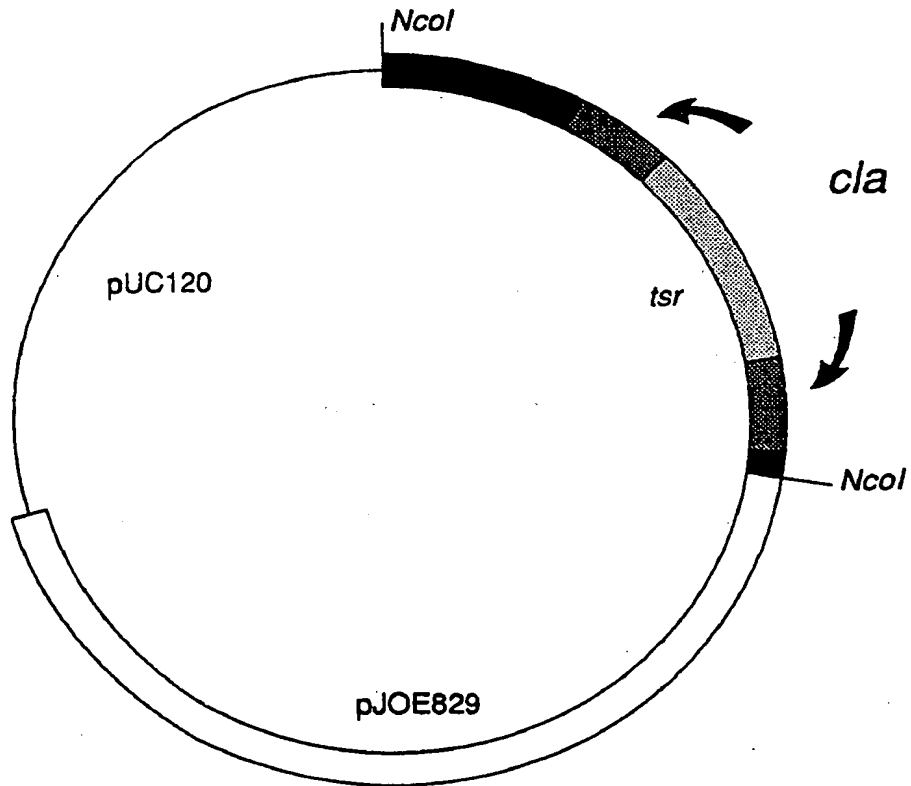


FIGURE 5

Sim; of. Barry

2108113

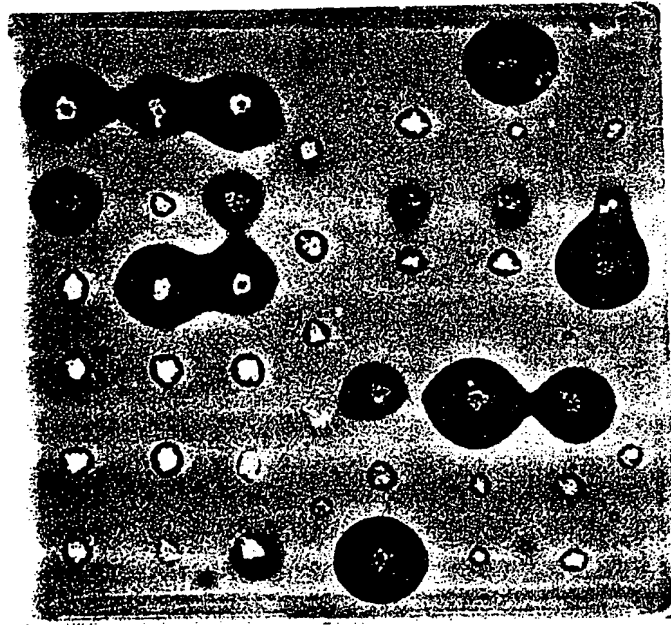


FIGURE 6

Sim, M. Bunn

60
 S. Cl. CLA 1 veridshvspryaqiptFmRLPhdpQPrGYDV--VvIGaPyDggTSyRpGARfGPqAIR
 E. co. AUH MSTIGHqYdNslvSnafGFIrLPmnfQPydsDadwVItGvPfDmaTSgRaGGRhGPaAIR
 yeast ARG MeT-GphY-NyyKnReIsIvIAPFSgGQgkIGVEKGPKymIKhGL-qtsIedlgwsteLE
 rat ARG MS-----sKpkpleIIGAPFSKGQPRGGVEKGPaLRKAGL-----vE
 human ARG MS-----aKSrtIGIIGAPFSKGQPRGGVEeGPTvLRKAGL-----LE

120
 S. Cl. CLA 61 seSgIihgvgIdRgPgtFDI---INcVdaGDINItPfDmniaIdtaQsHISgLLKANaaf
 E. co. AUH qvStnl-awehnRfPwnFDmrerINvVdcGDIVyafgDarEmSEkLQAHaeKLLaAGkrm
 yeast ARG psmdea-qfVgKikmekdsttgssVmidGVKakRadIVGEAtkIvynsVSKVvqANRfp
 rat ARG KLKEtE-ynV-rDhGDLafVdVpNDSPFQIVKNPRS--VGKAEQLAAvVAetaKNGtIS
 human ARG KLKEqE-cdV-KOyGDLpFadIPNDSPFQIVKNPRS--VGKASEQLagkVAqvKKNGRIS

180
 S. Cl. CLA 121 LmiGGDHSLTvaalRAVAeqhGpLAVVHIDAHsDTNpafyGgryhHGTPFrhgIdEkLI
 E. co. AUH LsfGGDHfvTIpILRAhAkhfGkmALVHfDAHTDTyan--GcefdHGTmFytapkeGLID
 yeast ARG LtLGGDHSAIGtVSAVIdkyPDaGLIWIDAHaDINTi--esTpSGNLHGcPVSFLmgIn
 rat ARG vVLGGDHsmaIGSISsHARVHPDLcVWVDAHTDINTP--LTTsSGNLHGQPVaFLLKEL
 human ARG LVLGGDHSLAIGSISgHARVHPDLGVWVDAHTDINTP--LTTtSGNLHGQPVVSFLKEL

240
 S. Cl. CLA 181 PaamVQIGIRGHNPkPDSLdyarghGvrVvtAdefgelgVggtadLirekV-----
 E. co. AUH PnhsvQIGIRt-----efdkdnGftVIdAcavnDrsVddvIaqvkaIV-----
 yeast ARG KdvphcpesIk-----WVpgnISpKkIaYIGLRDvDaGEkkILKdLGlaaFSMyhVD
 rat ARG KGKfPDVPGFS-----WVTPCISAKDIVYIGLRDvDPGEHYIITLGIKYFSMTEVD
 human ARG KGKIPDVPgFS-----WVTPCISAKDIVYIGLRDvDPGEHYIITLGIKYFSMTEVD

300
 S. Cl. CLA 241 -----GaRPVYvSvDIDvVDPaFAPGTGTPapGGLISREvLaLIR
 E. co. AUH -----GdmPVYLtFDIDcLDPaFAPGTGTPVIGGLTSdraikLYR
 yeast ARG KyGInaVIEamkavhpetnGegPImcSyDVGVDPIyIPATGTPVRGGLTIREGLfLVE
 rat ARG KLGIgKvME--ETfSYLLGRKKRPIHLSFDVDGLDPvFTPATGTPVVGGLsYREGLYITE
 human ARG rLGIGKvME--ETISYLLGRKKRPIHLSFDVDGLDPsFTPATGTPVVGGLTYREGLYITE

360
 S. Cl. CLA 301 cv-gDLkpVGfDVMEVsPIYDhggITsI-----IATeIgaELLYqyArahrTqIz
 E. co. AUH gL-KDLNIVGmDVVEVaPaYDaseITaI-----AAAtIALEmLYIqAaKkge
 yeast ARG rLaesGNLlaLDVVEcNPdLaihdIhVsnTisagcAIARcALGetII
 rat ARG EIYKTGLLSGLDIMEVNPTLGKTPEEVTRTVNTAVaITLsCFGtKREGNHKPeTDYLkPPK
 human ARG EIYKTGLLSGLDIMEVNPTLGKTPEEVTRTVNTAVaITLsCFGLaREGNHKPeIDYLnPPK

FIGURE 7

Simon M. Luning

2108113

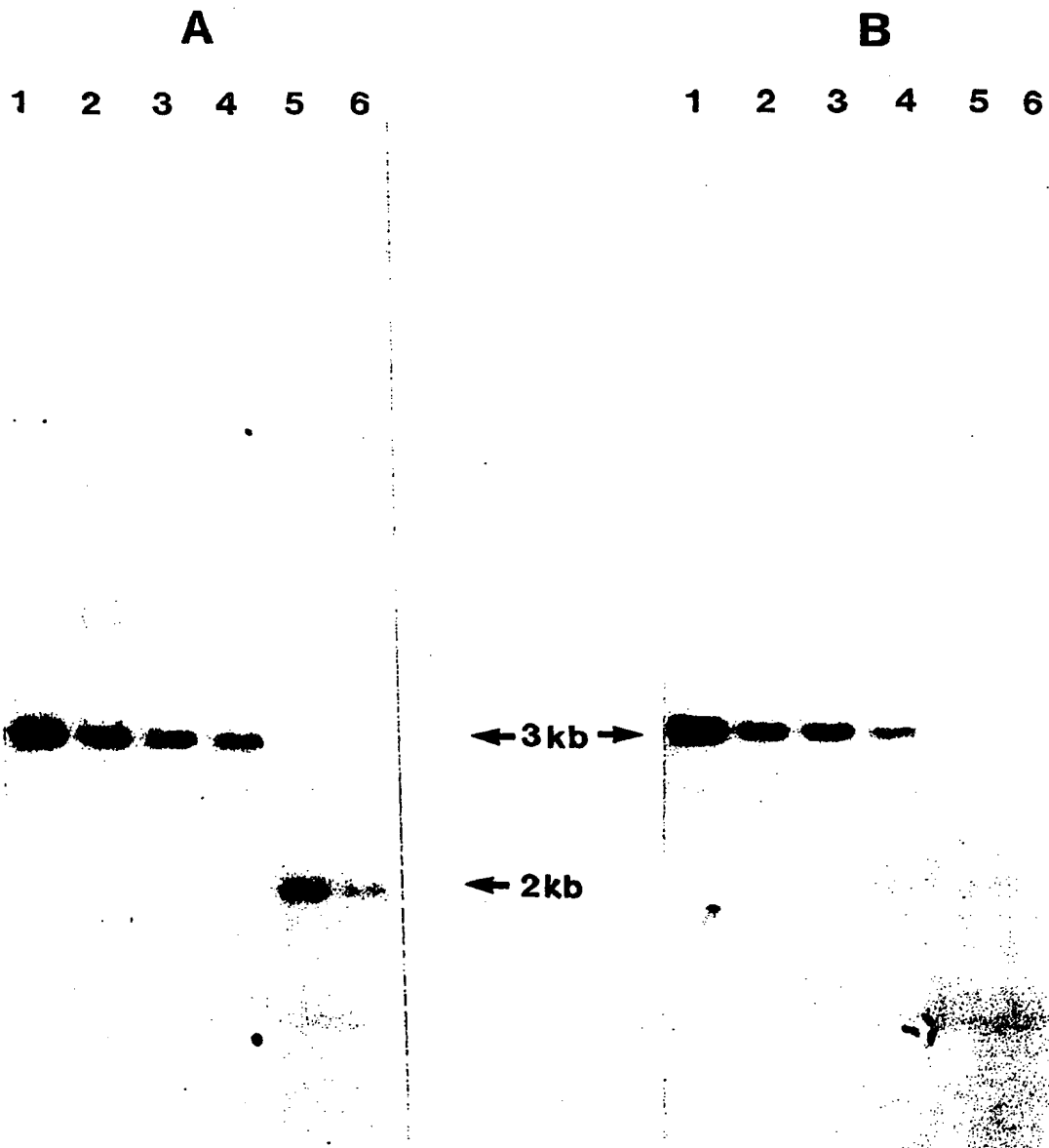


FIGURE 8

Sim; M. Bunn

Simon M. Baumgardner



2108113

	10	20	30	40	50	60	
1	MTHSDNYGDD	PPQGRRRSRG	RAATAVVAGL	AVTVGLGYWG	YTSLVADKED	SGDPEVEAAA	60
61	GQFDTFLGAW	EKGDAPTAAG	LTDTPDNAES	LIKSVMTNLK	PTKTEITAKT	GEKNPEGEVE	120
121	IPFTVRMTLP	GAGEYAWDST	AKVVGGGKEW	KVAFNTEMIH	PQMVPQOTLA	LKSRRERADIL	180
181	DANGNVLQAA	SIIGAVDPRT	GKGSAGLQSR	YDKQLTGGSG	AARSVVILDR	ESGOVVKKLT	240
241	GLKDTEGKPV	KTIDPRVQS	AAAAALEGSK	KNAATVAVDP	ATGNILAAAN	VPSGMNRALE	300
301	GRYPPGSTFK	VVTTAALLQQ	GMNPEERADC	PKFAHVNGQS	FENQDQFTLP	AGSTFRDSFA	360
361	HSCNTFFVNS	RSKLSESSLK	QAAEAFGIGG	TWDVGASTFD	GSPVVSNSEN	DKAASTIGQA	420
421	RVEASPLVMA	SIAATVKQGE	FKQFVLVPDA	VKKPHQAPRM	APGIVDSLRS	MMRSTVTDGA	480
481	GDALRGLGGQ	PHAKTGTAEF	GTEKPPKTHA	WMIGYQGDRN	IAWSVLLEDG	GSGGADAGPV	540
541	AAKFLSNLAA	GZ					552
	10	20	30	40	50	60	

FIGURE 10

Simon M. Burney

2108113

	10	20	30	40	50	60	
1	MSRVSTAPSG	KPTAAHALLS	RLRDHGVGKV	FGVVGREAAS	ILFDEVDPID	FVLTRHEFTA	60
61	GVAADVLARI	TGRPOACWAT	LPGGMINLST	GIATSVLDRS	PVIALAAQSE	SHDIFPNDTH	120
121	QCLDSVAIVA	PMSLYAVELQ	RPHEITDLVD	SAVNAAMTEP	VGPSFISLPV	DLLGSSEGID	180
181	TTVPNPPANT	PAKPVGVVAD	GWQKAADQAA	ALLAEAKHPV	LVVGAAAIRS	GAVPAIRALA	240
241	ERLNIPVITT	YIAKGVLPVG	HELVNYGAVTG	YMDGILNFPA	LQTMFAPVDL	VLTVGYDYAE	300
301	DLRPSMWQKG	IEKKTVRISP	TVNPIPRVYR	PDVDVVTDL	AFVEHFETAT	ASFGAKQRHD	360
361	IEPLRARIAE	FLADPETYED	GMRVHQVIDS	MNTVMEEAAE	PGEPTIVSDI	GFFRHYGVLF	420
421	ARADQPFGL	TSAGCSSFGY	GIPAAIGAQM	ARPDQPTFLI	AGDGGFHSNS	SDLETIARLN	480
481	LPIVTVVNN	DTNGLIELYQ	NIGHHRSHDP	AVKFGGVDFV	ALAEANGVDA	TRATNREELL	540
541	AALRKGAELG	RPFLIEVPVN	YDFQPGGFGA	LSIZ			574
	10	20	30	40	50	60	

FIGURE 11

Simon M. Lanning

2108113

	10	20	30		40	50	60	
1	MGAPVLPAAF	GFLASARTGG	GRAPGPFVAT	RGSHTDIDTP	QGERSLAATL	VHAPSVAPDR	60	
61	AVARSLTGAP	TTAVLAGEIY	NRDELLSVLP	AGPAPEGDAE	LVLRLLELYD	LHAFRLVNGR	120	
121	FATVVRTGDR	VLLATDHAGS	VPLYTCVAPG	EVRASTEAKA	LAHRDPKGF	PLADARRVAG	180	
181	LTGVYQVPAG	AVMDIDLGS	TAVTHRTWTP	GLSRRILPEG	EAVAAVRAAL	EKAVAQRVTP	240	
241	GDTPLVVLGS	GIDSSGVAAC	AHRAAGELDT	VSMGIDTSNE	FREARAVVDH	LRTRHREITI	300	
301	PTTELLAQLP	YAVWASESD	PDIEYLLPL	TALYRALDGP	ERRILTGYGA	DIPLGGMHRE	360	
361	DRLPALDTVL	AHDMATFDGL	NEMSPVLSTL	AGHWTHPYW	DREVLDLLVS	LEAGLKRRHG	420	
421	RDKWVLRAAM	ADALPAETVN	RPKLGVEHGS	GTTSSFSRLL	LDHGVAEDRV	HEAKQVVRE	480	
481	LFDLTVGGGR	HPSEVDIDDV	VRSVADRTAR	GAAZ			514	
	10	20	30	40	50	60		

Figure 12

Simon, J. Barry

2108113

	10	20	30	40	50	60	
1	VERIDSHVSP	RYAQIPTFMR	LPHDPQPRGY	DVVVIGAPYD	GGTSYRPGAR	FGPOAIRSES	60
61	GLINGVGIDR	GPGTFDLINC	VDAGDINLTP	FDMNIAIDTA	QSHLSGLLKA	NAAFLMIGGD	120
121	HSLTVAALRA	VAEQHGFLAV	VHLDHSDTN	PAFYGGRYHH	GTPFRHGIDE	KLIDPAAMVQ	180
181	IGIRGHNPKE	DSL DYARGHG	VRVVTADDFG	ELGVGGTADL	IREKVGQRPV	YVSVDIDVVD	240
241	PAFAPGTGTP	APGGLLSREV	LALLRCVGDL	KPVGFDVMEV	SPLYDHGGIT	SILATEIGAE	300
301	LLYQYARAHR	TQLZ					314
	10	20	30	40	50	60	

FIGURE 13

Simon, M. Barry

2108113

	10	20	30	40	50	60	
1	MASPIVDCTP	YRDELLALAS	ELPEVPRADL	HGFLDEAKTL	AARLPEGLAA	ALDTFNAVGS	60
61	EDGYLLLRGL	PVDDSELPET	PTSTPAPLDR	KRLVMEAMLA	LAGRRLGLHT	GYQELRSGTV	120
121	YHDVYPSPGA	HYLSSETSET	LLEFHTEMAY	HILQPNYVML	ACSRADHENR	AETLVGSVRK	180
181	ALPLLDEKTR	ARLFDRKVPC	CVDVAFRGGV	DDPGALANVK	PLYGDANDPF	LGYDRELLAP	240
241	EDPADKEAVA	HLSQALDDVT	VGVKLVPGDV	LIIDNFRTH	ARTPFSPRWD	GKDRWLHRVY	300
301	IRTDNRNGQLS	GGERAGDTIS	FSPRRZ				326
	10	20	30	40	50	60	

FIGURE 14

Simon; M. Lunny

2108113

	10	20	30	40	50	60	
1	MSDSTPKTPR	GFVVHTAPVG	LADDGRHDFT	VLASTAPATV	SAVFTRSRFA	GPSVVLCREA	60
61	VADGQARGVV	VLARNANVAT	GLEGEENARE	VREAVARALG	LPEGEMLIAS	TGVIGRQYPM	120
121	ESIREHLKTL	EWPAGEGGFD	RAARAIMTTD	TRPKEVRVSV	GGATLVGLAK	GVGMLEPDMA	180
181	TLLTFFATDA	RLDPAEQDRL	FRRVMDRTFN	AVSIDTDTST	SDTAVLFANG	LAGEVDAGEF	240
241	EEALHTAALA	LVKDIASDGE	GAAKLIEVQV	TGARDDAQAK	RVGKTIVVNSP	LVKTAVHGCD	300
301	PNWGRVAMAI	GKCSDDTDID	QERVITIRFE	VEVYPPKARG	DQADDALRAA	VAEHLRGDEV	360
361	VIGIDLAIAD	GAFTVYGCDL	TEGYVRLNSE	YTTZ			394
	10	20	30	40	50	60	

FIGURE 15

Sim; M. Barry

2108113

		10		20		30		40		50		60	
1		METTRSTTAD		EGFDAGVRGV		VAPTDAPGGT		LRLVRTDDFD		SLDPGNTYYA		YTNWFLRLIG	60
61		RTLVTFTDAP		GKAGQRLVPD		LAESLGESSE		DGRVWYRLR		EGLRYEDGTP		VVSADIKHAI	120
121		ARSNYGTDVL		GAGPTYFRHL		LGTEYGGPWR		EPDADGPVTL		ETPDERTLVF		RLREPFAGMD	180
181		LLATMPSTTP		VPRDRDTGAE		YRLRPVATGP		YRIVSYTRGE		LAVLEPNPHW		DPETDPVRVQ	240
241		RASRIEVHLG		KDPHEVDRML		LAGEAHVDLA		GFGVQPAAQE		RILAEPELRA		HADNPLTGFT	300
301		WIYCLSSRIA		PFDNVHCRR		VQFATDKAAM		QEAYGGAVGG		DIATTLLPPT		LDGYKHFDY	360
361		PVGPEGTGDL		EAARAEKLA		GMPDGFRTRI		AARKDRLKEY		RAAEALAAGL		ARVGIEAEVL	420
421		DFPSGDYFDR		YGGCPEYLRE		HGIGIMFGW		GADFPDGYGF		LQIITDGRAI		KERGNQNMGE	480
481		LDDPEINALL		DEGAQCADPA		RRAEIWHRID		QLTMDHAVIV		PYLYPRSLLY		RHPDTRNAFV	540
541		TGSFGMYDYV		ALGAKZ									556
		10		20		30		40		50		60	

FIGURE 16

Sim; M. Barney

2108113

	10	20	30	40	50	60	
1	MEVARRTGVR	HGTVERRLDR	LDRIVGLPLT	LRSRHTARLT	TAGSRILVAG	RRFFHQVDLA	60
61	ARTHIFGHGS	EAVDAPEVLS	LVSTEPLLDE	VVEDAAASLD	LLLSVRHEAP	HQVAAQLAGY	120
121	QVDAAYTWSL	QSPRHSLERS	VRTCEVLDDP	LWVILPRDHP	LAARREVSLA	DLRDET WVSE	180
181	TGPGSEILVT	RVFQLAGLTA	PTRLHITGAS	VARGILRRGD	AIGLGSPTHP	AVQDPSLVRR	240
241	SLAERPRRTT	SLLVDPTIVP	RALAGRLAAL	IAEVQLRRFA	EHHRDLLDEP	WWAQWYAERT	300
301	GADARRFGAG	PDQGSVPGQA	EGRKLDVDDL	HLLQAVARHG	SINRAAAVLS	ISQSALTRRI	360
361	HRLEQSLGAR	LLLRSPRGTS	LTGPTRQFLR	QLALYEAEFR	EAALACRSVE	RPLAQGHWPI	420
421	RRGVAAGARM	SGZ					433
	10	20	30	40	50	60	

FIGURE 17

Sim; M. Barry

2108113

	10	20	30	40	50	60	
1	MPSALQKVA	LITGASSGIG	EATARALAAE	GAAVAIAARR	VEKLRALGDE	LTAAGAKVHV	60
61	LELDVADROG	VDAAVASTVE	ALGGLDILVN	NAGIMLLGPV	EDADTTDWTR	MIDTNLLGLM	120
121	YMTRAALPHL	LRSKGTVVQM	SSIAGRVNVR	NAAVYQATKF	GVNAFSETLR	QEVTERGVRV	180
181	VVIEPGTTDT	ELRGHITHTA	TKEMYEQRIS	QIRKLQAQDI	AEAVRYAVTA	PHHATVHEIF	240
241	IRPTDQVZ						248
	10	20	30	40	50	60	

FIGURE 18

Sim; M. Barry

	10	20	30	40	50	60	
1	MMNEAAPQSD	QVAPAYPMHR	VCPVDPPPQL	AGLRSQKAAS	RVTLWDGSQV	WLVTSAGAR	60
61	AVLGDRRFTA	VTSAPGFPML	TRTSQVLVRAN	PESASFIRMD	DPQHSRLRSM	LTRDFLARRA	120
121	EALRPVREL	LDEILGGLVK	GERPVDLVAG	LTIPVPSRVI	TLLFGAGDDR	REFIEDRSV	180
181	LIDRGYTPEQ	VAKARDEL DG	YLRELVEERI	ENPGTDLISR	LVIDQVRPGH	LRVEEMVPMC	240
241	RLLLVAGHGT	TTSQASLSLL	SLLTOPELAG	RLTEDPALLP	KAVEELLRFH	SIVQNGLARA	300
301	AVEDVQLDDV	LIRAGEGVVL	SLSAGNRDET	VFPDPDRVDV	DRDARRHLAF	GHGMHQCLGQ	360
361	WLARVELEEI	LA AVL RWMPG	ARLAVPFEEL	DFRHEVSSYG	LGALPVTWZ		409
	10	20	30	40	50	60	

FIGURE 19

Simon, M. Lunny

THIS PAGE BLANK (USPTO)

**This Page is Inserted by IFW Indexing and Scanning
Operations and is not part of the Official Record**

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☒ **BLACK BORDERS**
- ☐ **IMAGE CUT OFF AT TOP, BOTTOM OR SIDES**
- ☐ **FADED TEXT OR DRAWING**
- ☐ **BLURRED OR ILLEGIBLE TEXT OR DRAWING**
- ☐ **SKEWED/SLANTED IMAGES**
- ☐ **COLOR OR BLACK AND WHITE PHOTOGRAPHS**
- ☐ **GRAY SCALE DOCUMENTS**
- ☐ **LINES OR MARKS ON ORIGINAL DOCUMENT**
- ☐ **REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY**
- ☐ **OTHER:** _____

IMAGES ARE BEST AVAILABLE COPY.

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.

THIS PAGE BLANK (USPTO)